

SEQUENCE LISTING

<110> Glucksmann, Maria A.
Hodge, Martin G.

#4

<120> 2871 RECEPTOR, A NOVEL G-PROTEIN COUPLED RECEPTOR

<130> 5800-2A (035800/183295)

<140> US 09/324,465

<141> 1999-06-02

<150> 09/088,857

<151> 1998-06-02

<160> 6

<170> PatentIn Ver. 2.0

<210> 1

<211> 358

<212> PRT

<213> Homo sapiens

<400> 1

Met Gly Phe Asn Leu Thr Leu Ala Lys Leu Pro Asn Asn Glu Leu His
1 5 10 15

Gly Gln Glu Ser His Asn Ser Gly Asn Arg Ser Asp Gly Pro Gly Lys
20 25 30

Asn Thr Thr Leu His Asn Glu Phe Asp Thr Ile Val Leu Pro Val Leu
35 40 45

Tyr Leu Ile Ile Phe Val Ala Ser Ile Leu Leu Asn Gly Leu Ala Val
50 55 60

Trp Ile Phe Phe His Ile Arg Asn Lys Thr Ser Phe Ile Phe Tyr Leu
65 70 75 80

Lys Asn Ile Val Val Ala Asp Leu Ile Met Thr Leu Thr Phe Pro Phe
85 90 95

Arg Ile Val His Asp Ala Gly Phe Gly Pro Trp Tyr Phe Lys Phe Ile
100 105 110

Leu Cys Arg Tyr Thr Ser Val Leu Phe Tyr Ala Asn Met Tyr Thr Ser
115 120 125

Ile Val Phe Leu Gly Leu Ile Ser Ile Asp Arg Tyr Leu Lys Val Val
130 135 140

Lys Pro Phe Gly Asp Ser Arg Met Tyr Ser Ile Thr Phe Thr Lys Val
145 150 155 160

Leu Ser Val Cys Val Trp Val Ile Met Ala Val Leu Ser Leu Pro Asn
165 170 175

Ile Ile Leu Thr Asn Gly Gln Pro Thr Glu Asp Asn Ile His Asp Cys
180 185 190

Ser Lys Leu Lys Ser Pro Leu Gly Val Lys Trp His Thr Ala Val Thr
195 200 205

Tyr Val Asn Ser Cys Leu Phe Val Ala Val Leu Val Ile Leu Ile Gly
210 215 220

Cys Tyr Ile Ala Ile Ser Arg Tyr Ile His Lys Ser Ser Arg Gln Phe
225 230 235 240

Ile Ser Gln Ser Ser Arg Lys Arg Lys His Asn Gln Ser Ile Arg Val
245 250 255

Val Val Ala Val Phe Phe Thr Cys Phe Leu Pro Tyr His Leu Cys Arg
260 265 270

Ile Pro Phe Thr Phe Ser His Leu Asp Arg Leu Leu Asp Glu Ser Ala
275 280 285

Gln Lys Ile Leu Tyr Tyr Cys Lys Glu Ile Thr Leu Phe Leu Ser Ala
290 295 300

Cys Asn Val Cys Leu Asp Pro Ile Ile Tyr Phe Phe Met Cys Arg Ser
305 310 315 320

Phe Ser Arg Arg Leu Phe Lys Lys Ser Asn Ile Arg Thr Arg Ser Glu
325 330 335

Ser Ile Arg Ser Leu Gln Ser Val Arg Arg Ser Glu Val Arg Ile Tyr
340 345 350

Tyr Asp Tyr Thr Asp Val
355

<210> 2

<211> 1489
 <212> DNA
 <213> Homo sapiens

<400> 2
 ccacgcgtcc ggagaatttg aaaggggtgcc ccaaaggaca atctctaaag gggtaagggg 60
 gatacctacc ttgtctggtg ggggagatgt ttcgttttca tgctttacca gaaaatccac 120
 ttccctgccc accttagttt caaagcttat tcttaattag agacaagaaa cctgtttcaa 180
 cttgaagaca ccgtatgagg tgaatggaca gccagccacc acaatgaaag aaatcaaacc 240
 aggaataacc tatgctgaac ccacgcctca atcgccccca agtgtttcct gacacgcac 300
 tttgcttaca gtgcatcaca actgaagaat ggggttcaac ttgacgcttg caaaattacc 360
 aaataacgag ctgcacggcc aagagagtca caattcaggc aacaggagcg acgggccagg 420
 aaagaacacc acccttcaca atgaatttga cacaattgtc ttgccggtgc tttatctcat 480
 tatatttgtg gcaagcatct tgctgaatgg tttagcagtg tggatcttct tccacattag 540
 gaataaaacc agcttcata tctatctcaa aaacatagtg gttgcagacc tcataatgac 600
 gctgacattt ccatttcgaa tagtccatga tgcaggattt ggaccttggg acttcaagtt 660
 tattctctgc agatacactt cagttttgtt ttatgcaaac atgtatactt ccatcgtgtt 720
 ccttgggctg ataagcattg atcgctatct gaagggtgtc aagccatttg gggactctcg 780
 gatgtacagc ataaccttca cgaaggtttt atctgtttgt gtttgggtga tcatggctgt 840
 tttgtctttg ccaaaccatca tcttgacaaa tggtcagcca acagaggaca atatccatga 900
 ctgctcaaaa cttaaaagtc ctttgggggt caaatggcat acggcagtc cctatgtgaa 960
 cagctgcttg tttgtggccg tgctgggtgat tctgatcgga tgttacatag ccatatccag 1020
 gtacatccac aaatccagca ggcaattcat aagtcagtc agccgaaagc gaaaacataa 1080
 ccagagcatc aggggttgtt tggtctgtgt ttttacctgc tttctaccat atcacttgtg 1140
 cagaattcct tttactttta gtcacttaga caggctttta gatgaatctg cacaaaaaat 1200
 cctatattac tgcaaagaaa ttacactttt cttgtctgcg tgtaatgttt gcctggatcc 1260
 aataatttac tttttcatgt gtaggtcatt ttcaagaagg ctgttcaaaa aatcaaatat 1320
 cagaaccagg agtgaagca tcagatcact gcaaagtgtg agaagatcgg aagttcgcac 1380
 atattatgat tacactgatg tgtaggcctt ttattgtttg ttggaatcga tatgtacaaa 1440
 gtgtaaataa atgtttcttt tcattaataa aamaaaaaaa aaaaaaaag 1489

<210> 3
 <211> 269
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: consensus
 sequence of the seven transmembrane domain
 rhodopsin superfamily from the Prosite data base

<400> 3
 Gly Asn Ile Leu Val Ile Trp Val Ile Cys Arg Tyr Arg Arg Met Arg
 1 5 10 15
 Thr Pro Met Asn Tyr Phe Ile Val Asn Leu Ala Val Ala Asp Leu Leu
 20 25 30

Phe Ser Leu Phe Thr Met Pro Phe Trp Met Val Tyr Tyr Val Met Gln
35 40 45
Gly Arg Trp Pro Phe Gly Asp Phe Met Cys Arg Ile Trp Met Tyr Phe
50 55 60
Asp Tyr Met Asn Met Tyr Ala Ser Ile Phe Phe Leu Thr Cys Ile Ser
65 70 75 80
Ile Asp Arg Tyr Leu Trp Ala Ile Cys His Pro Met Arg Tyr Met Arg
85 90 95
Trp Met Thr Pro Arg His Arg Ala Trp Val Met Ile Ile Ile Ile Trp
100 105 110
Val Met Ser Phe Leu Ile Ser Met Pro Pro Phe Leu Met Phe Arg Trp
115 120 125
Ser Thr Tyr Arg Asp Glu Asn Glu Trp Asn Met Thr Trp Cys Met Ile
130 135 140
Tyr Asp Trp Pro Glu Trp Met Trp Arg Trp Tyr Val Ile Leu Met Thr
145 150 155 160
Ile Ile Met Gly Phe Tyr Ile Pro Met Ile Ile Met Leu Phe Cys Tyr
165 170 175
Trp Arg Ile Tyr Arg Ile Ala Arg Leu Trp Met Arg Met Ile Pro Ser
180 185 190
Trp Gln Arg Arg Arg Arg Met Ser Met Arg Arg Glu Arg Arg Ile Val
195 200 205
Lys Met Leu Ile Ile Ile Met Val Val Phe Ile Ile Cys Trp Leu Pro
210 215 220
Tyr Phe Ile Val Met Phe Met Asp Thr Leu Met Met Trp Trp Phe Cys
225 230 235 240
Glu Phe Cys Ile Trp Arg Arg Leu Trp Met Tyr Ile Phe Glu Trp Leu
245 250 255
Ala Tyr Val Asn Cys Pro Cys Ile Asn Pro Ile Ile Tyr
260 265

<210> 4
<211> 20

<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide primer

<400> 4
atcgtgttcc ttgggctgat

20

<210> 5

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide primer

<400> 5
tccgagagtc cccaaatgg

19

<210> 6

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide probe

<400> 6
agcattgatc gctatctgaa ggtggtcaa

29

CCAAGGTTCCGACAAATTTGAAAGGGTGGCCCAAGGACAAATCTCTAAAGGGTAAGGGAGATACCTACCTTGTCTGGT
 AAGGGAGATGTTTCTGTTTTCATGCTTTTACCAGAAAATGCACTTCCCTCCGAGCTTAGTTTCAAGCTTATTCTTAATT
 AGAGACAAGAAAACCTGTTTCAACTTGAAGACACCGTATGAGGTGATGGAACCCAGCCACCACAAATGAAGAAATCAA
 ACCAGCAATTAACCTATGCTGAAACCCAGGCTCAATGGTGGCCAGTCTTTCTGACAGCATCTTTGCTTACAGTGCAT
 H G F N L T L A K L P N N E L H 16
 CACAACTGAAGA ATG GGG TTC AAC TTG ACG CTT GCA AAA TTA CCA AAT AAC GAG CTG CAC 48
 G Q E S H N S G N R S D G P G K N T T L 36
 GGC CAA GAG AGT CAC AAT TCA GGC AAC AGG AGC GAG GGG CCA CGA AAG AAC ACC ACC CTT 108
 H N E F D T I V L P V L Y L I I F V A S 56
 CAC AAT GAA TTT GAC ACA ATT GTC TTG CCG GTG CTT TAT CTC AAT ATA TTT CTG GCA AGC 168
 I L L N G L A V W I F P H I R N K T S P 76
 ATC TTG CTG AAT GGT TTA GCA GTG TGG ATC TTC TTC CAC ATT AGG AAT AAA ACC AGC TTC 228
 I F Y L K N I V V A D L I M T L T F P F 96
 ATA TTC TAT CTC AAA AAC ATA GTC GTT CCA GAC CTC ATA ATG ACG CTC ACA TTT CCA TTT 288
 R I V H D A G F G P W Y F K F I L C R Y 116
 CGA ATA GTC CAT GAT CCA GGA TTT GGA CCT TGG TAC TTC AAG TTT ATT CTC TGC AGA TAC 348
 T S V L F Y A N H Y T S I V F L G L Y S 136
 ACT TCA GTT TTG TTT TAT CCA AAC ATG TAT ACT TCC ATC GTG TTC CTT CCG CTG ATA AGC 408
 I D R Y L K V V K P F G D S R H Y S I T 156
 ATG GAT CCG TAT CTG AAG GTG GTC AAC CCA TTT GGG GAC TCT CCG ATG TAC AGC ATA ACC 468
 F T K V L S V C V W V I H A V L S L P N 176
 TTC ACG AAG GTT TTA TCT GTT TGT GTT TGG GTG ATC ATG GCT GTT TTG TCT TTG CCA AAG 528
 I I L T N G Q P T E D N I H D C S K L K 196
 ATC ATC CTG ACA AAT GGT CAG CCA ACA GAG GAC AAT ATC CAT GAC TGC TCA AAA CTT AAA 588
 S P L G V K W H T A V T Y V N S C L F V 216
 AGT CCT TTG CCG GTC AAA TGG CAT ACG CCA GTC ACG TAT GTG AAC AGC TGG TTG TTT GTG 648
 A V L V I I G C Y I A I S R Y I H K S 236
 CCC GTG CTG GTG ATT CTG ATC GGA TGT TAC ATA GCC ATA TCC AGG TAC ATC CAC AAA TCC 708
 S R Q F I S Q S S R K R K H N Q S I R V 256
 ACC AGG CAA TTC ATA AGT CAG TCA AGC CGA AAG CGA AAA CAT AAC CAG AGC ATC AGG GTT 768
 V V A V F F T C F L P Y H L C R I P F T 276
 GTT CTC GCT GTC TTT TTT ACC TGC TTT CTA CCA TAT CAC TTG TGC AGA ATT OCT TTT ACT 828
 F S H L D R L L D E S A Q K I L Y Y C K 296
 TTT AGT CAC TTA GAC ACG CTT TTA GAT GAA TCT GCA CAA AAA ATC CTA TAT TAC TGC AAA 888
 E I T L F L S A C N V C L D P I I Y F F 316
 GAA ATT ACA CTT TTC TTG TCT GCG TGT AAT GTT TGC CTG GAT CCA ATA ATT TAC TTT TTC 948
 H C R S F S R R L F K K S N I R T R S E 336
 ATG TGT AGG TCA TTT TCA AGA AGG CTG TTC AAA AAA TCA AAT ATC AGA ACC AGG AGT GAA 1008
 S I R S L Q S V R R S E V R I Y Y D Y T 356
 ACC ATC AGA TCA CTG CAA AGT GTG AAG AGA TCG GAA GTT CCG ATA TAT TAT GAT TAC ACT 1068
 D V
 GAT GTG TAG 359
 1077
 CCTTTTATGTTTGTTCGAAATCGATATGTACAAAGTGTAAATAAATGTTTCTTTTCATTAATAAAGAAAAA
 AAAAG

FIG 1

Score: 191.33 Seq: 59 314 Model: 1 255
 *GNILVIWVICRYRRMRTPMNYFIVNLAVADLLF=1fEMPFWMVYVVMGG
 N+L +W+++ R +T++ +++ N VAOL++ ++T+FF +V+ + G
 f1h2871 59 LNLAVWVIFPHRN-KTSFIFYLKNIVVADLIH-TLTFPFRIVHDAGFG 105
 RWpFGd+MCRiWmYFDYHMMVASIFELTcISIDRYLWAIChPN-yMRWMT
 W+S ++CR ++ ++Y NMY+SI FL +ISIDRYL ---R+ R+ -
 f1h2871 106 PHYFKFYLCRYTSVLFYANMYTSIVFLGLISIDRYL-KVVKPFQDSRYYS 154
 PRNRRAWVNI111YWMVSF1ISMPPFLMFrWscyrDeneWNgTWCMiYDWPa
 +++ +-V+++++WV+++++S+P + +r + + ++ E-N+ C+ - P
 f1h2871 158 ITP-TKVLSVCWVWVNAVLSLP-N-II-LTNGQ?-TEDNINOC5XKXSP 199
 .WMHWVWVILmciingFYIPNIIM1FCYWRIYRIaRIWMXNips-C+FR+
 + W - V--- - F- - I - CY I R +---R- S--
 f1h2871 200 GYXNHTAVCVVNS-CLFVAVLVLLGCVIAISRYIHXSBQFISGES--- 245
 m5MRrERRRIVXMI11HVVFFIICW1PYFIWmMDTLH.XXWFCeSC.IWf
 R+R- - --- VF+ C+LPY++ + T- - -
 f1h2871 246 ----RKRKHNSQIRVVVAVVFTCFILPYHLCAIPFQPSHLDRLLD-ESACK 290
 f1wa.Y.IfeWLaYVNCpCINSIIY*
 +L++ I+++L+ +N C++PIIY
 f1h2871 291 ILYYCKEITLFLSACIV-CLDPIIY 314

FIG 2

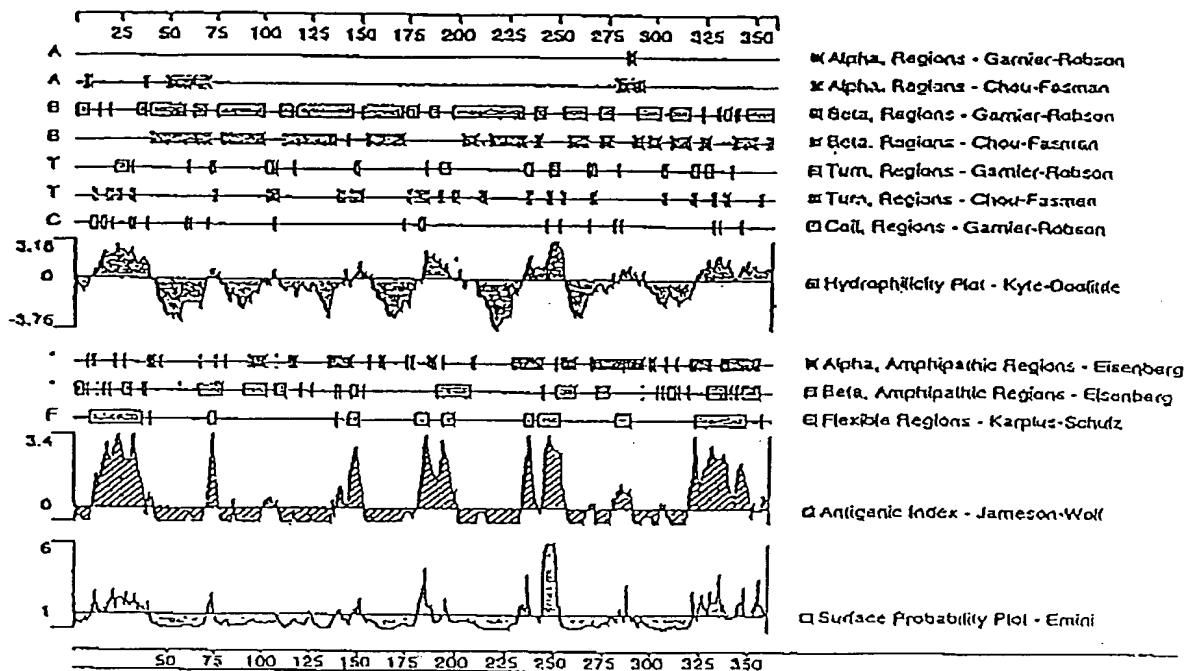


FIG 3

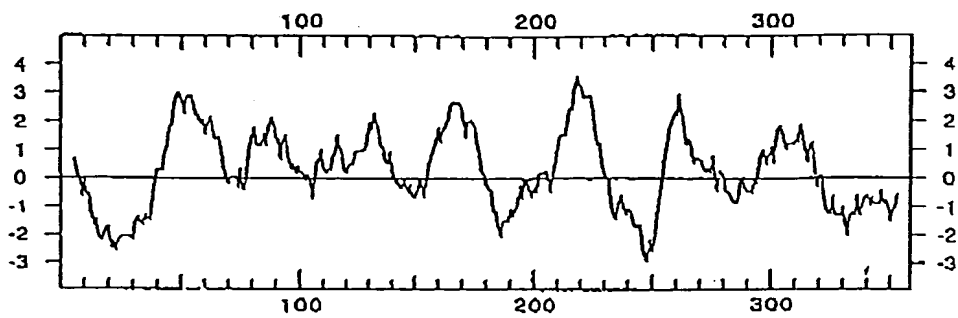
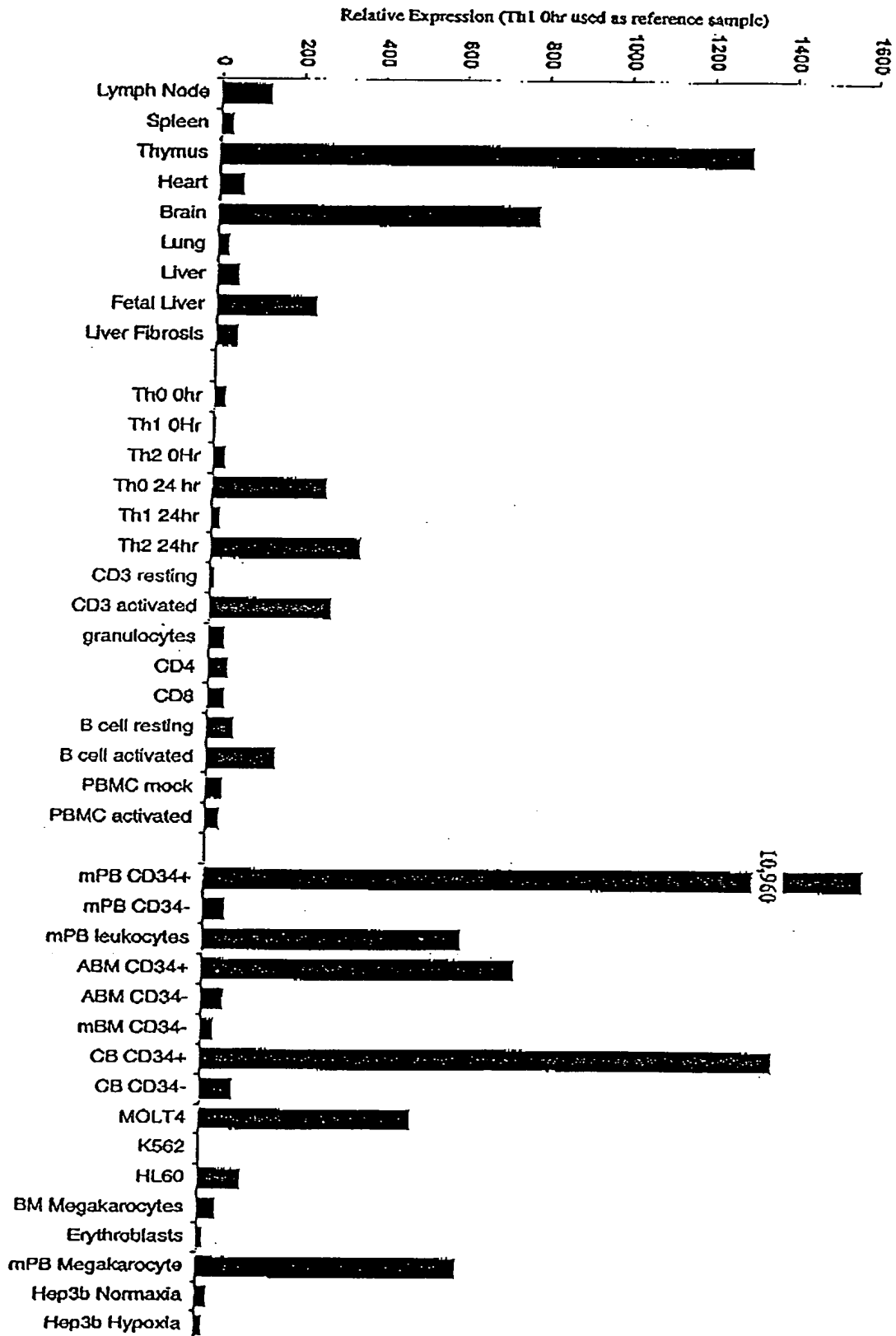


FIG 4

LEGAL DEPT.



2871 Expression

FIG 5

Relative Expression (Use colon as Reference)

1,000 2,000 3,000 4,000 5,000 6,000 7,000 8,000 9,000

Gene 2871 Expression in Normal Human Tissues

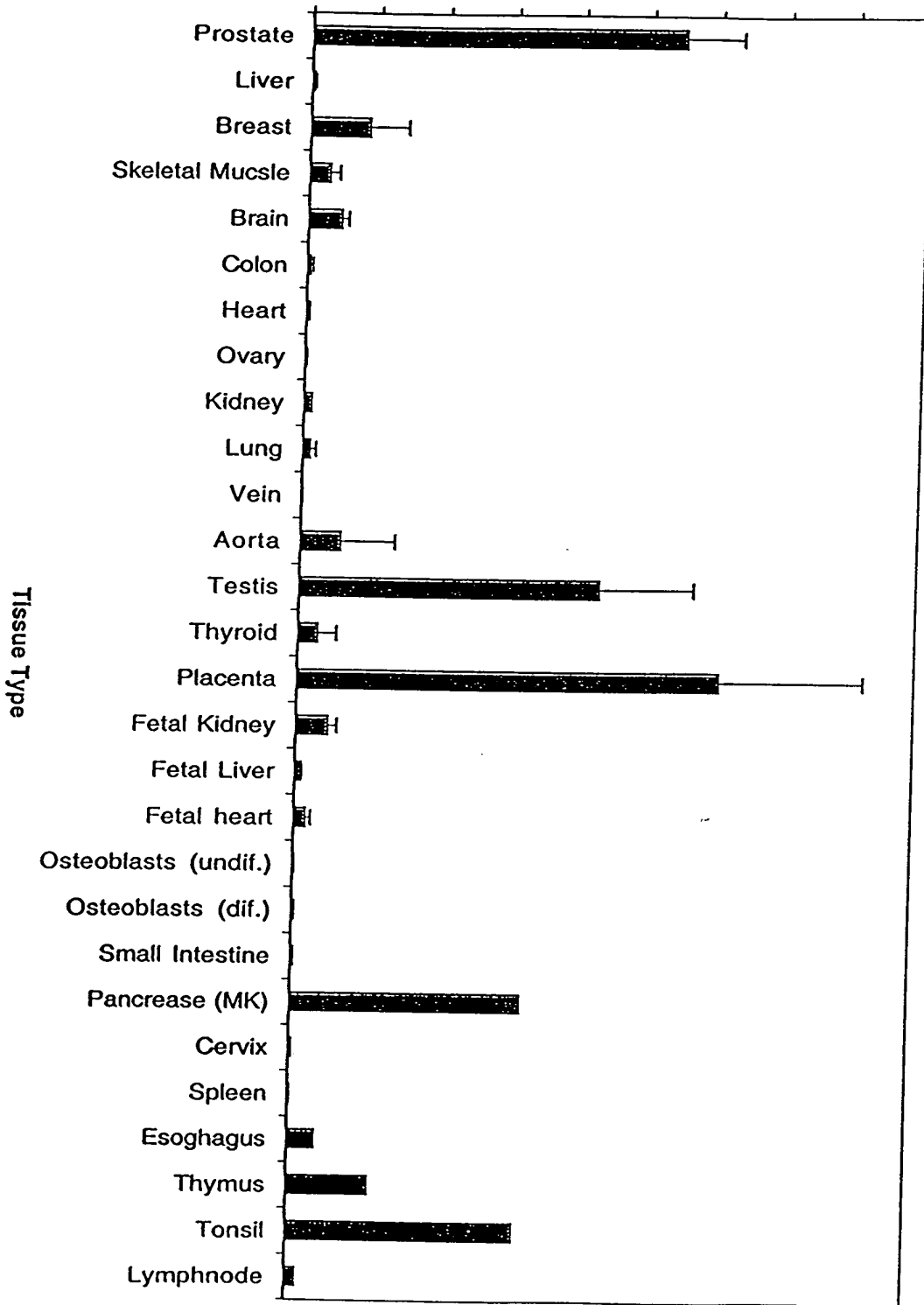


FIG 6

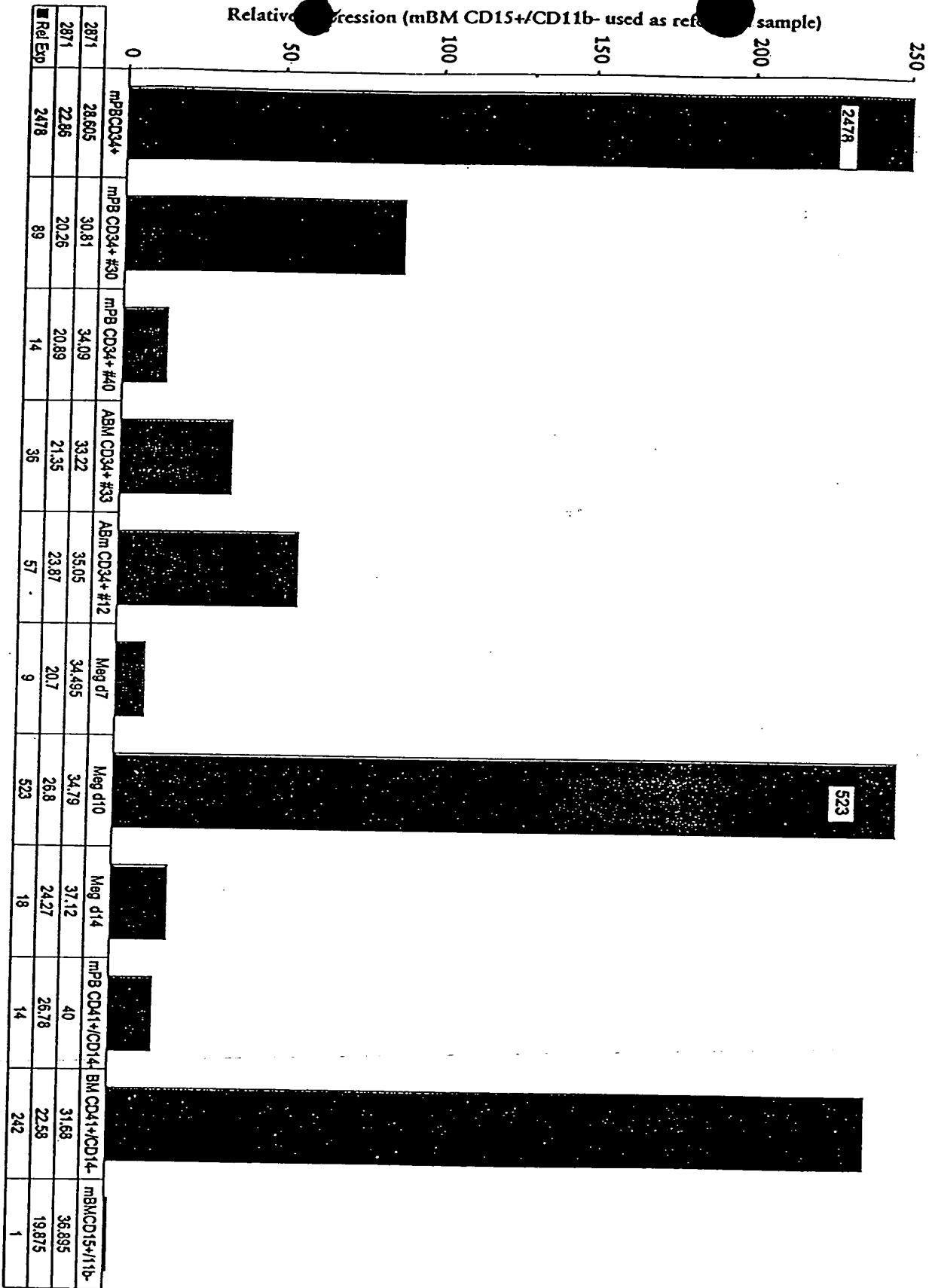


FIG 7